The geography of conflicts and regional trade agreements
Si Vis Pacem Para Mercatum[*]

Philippe Martin[†] Thierry Mayer[‡] Mathias Thoenig[§]

January 4, 2010

Abstract

We analyze the interaction of economic and political determinants of regional trade agreements (RTA). In addition to standard trade gains, RTAs provide peace-promoting benefits by offering a political forum and by increasing the opportunity cost of trade destroying wars. If policy makers believe in the pacifying effect of RTAs, country-pairs with large trade gains from RTAs **and** high probability of conflict are more likely to sign an RTA. Using data on the 1950-2000 period, we show that this complementarity between economic and political gains is at work in the geography of RTAs. Country pairs characterized by a high frequency of old wars (which we use as a proxy of the probability of conflict) are shown to be more likely to sign RTAs, the more so the higher the trade gains from an RTA. These trade gains are estimated by a theory-driven empirical strategy to disentangle them from the political factors. We also show that, contrary to old wars, recent wars make it more difficult to negotiate an RTA. This suggests the existence of windows of opportunity to lock-in RTAs and peace. Finally multilateral trade openness, because it increases the probability of a bilateral conflict, increases the political incentive to sign RTAs.

JEL classification: F12, F15 Keywords: Regional trading arrangements, Trade, War.

[*] "If you want peace, prepare for trade". This paper is part of the research project "Politics, Economics and Global Governance: The European Dimensions" (PEGGED), Seventh Research Framework Programme, Contract number 217559. Philippe Martin and Thierry Mayer thank Institut Universitaire de France, and ANR for financial assistance. We are grateful to Séverine Toussaert for her research assistance, and for participants at seminars in WTO, EUI,... for their comments. All errors remain ours.

[†] Sciences-Po and CEPR

[‡] Sciences-Po and CEPR

[§] University of Lausanne and CEPR

1 Introduction

Regional trade agreements (RTAs) have a bad press among a number of economists. Many scholars argue that they constitute a threat to the carefully constructed postwar multilateral trade system. Whereas multilateral trade liberalization has stalled, the number of RTAs has massively expanded during the last two decades and they are now well over 300. The well known problem with these bilateral and regional agreements is that, although they create trade, they also generate distortions by excluding countries.[1] Much less attention has been paid (by economists) to the political and strategic motivations for regional integration, even though these motivations may have been key historically.[2] In fact the debate between economists and political scientists often interprets economic and political rationales for RTAs as substitutes. In this paper, we revisit the case for regional integration by explicitly linking the economic and political rationales and we argue that both theoretically and empirically they are complement.

One key link is provided by the so called Liberal Peace argument which states that bilateral trade dependency reduces the probability of a bilateral war, a mechanism that has been analyzed theoretically and on which some empirical evidence exists [3]. The next logical step is that RTAs, because they create trade, reduce the probability of wars between countries. This proposition is however difficult to test because establishing the direction of causality is a hard task: RTAs may reduce conflictuality but peace (or expected peace) may facilitate RTA negotiations. The lack of historical perspective *following* RTA formation (most RTAs were signed in the 1990s and 2000s) also makes identification difficult in the panel dimension.[4] We choose a different route by asking the following question: is the geography of RTAs consistent with a model in which policy makers believe that RTAs are pacifying and therefore believe in the Liberal Peace argument? This empirical strategy allows to exploit the period *preceding* RTAs formation for identifying the relevant effects.

[1] The most recent evidence (Baier and Bergstrand, 2007 using gravity equations) on trade creation finds a relatively large effect: RTAs are on average responsible for a doubling of trade between two members after 10 years. Baier and Bergstrand (2008) use matching techniques confirm this large effect of RTAs on trade between members. Much less is known about trade diversion and therefore the potential economic costs of these preferential agreements. Hence, the economic case for RTAs is still an open debate.

[2] In the case of Europe, political scientists and historians have insisted on the fact that economic integration was viewed as an intermediate objective while its final objective was to prevent the killing and destruction of the two World Wars from ever happening again. Even the recent creation of the euro, often interpreted by economists as a logical step towards more economic integration, has been discussed in these terms. Indeed, Jacques Delors (former president of the European Commission) declared: "...people forget too often about the political objectives of the European constitution. The argument in favor of the single currency should be based on the desire to live together in peace". Before that, the 1860 Anglo-French commercial Treaty was signed to diffuse tensions between the two countries. Outside Europe, MERCOSUR was created in 1991 in part to curtail the military power in Argentina and Brazil, then two recent and fragile democracies with potential conflicts over natural resources.

[3] see Oneal and Russett, 1999, Polachek, 1980, Martin, Mayer and Thoenig, 2008, Spolaore and Wacziarg, 2009

[4] Mansfield and Pevehouse (2000) find that country pairs in RTAs are less likely to be in conflict than others. However, their cross-sectional evidence does not allow to conclude on the direction of causality.

We first use a simple theoretical framework to illustrate the different mechanisms at work in the decision whether to sign or not an RTA. In addition to standard trade gains, leaders consider that RTAs provide two types of peace-promoting security gains (i) by offering a political forum which facilitates settlement of future disputes; (ii) by increasing the opportunity cost of future and potentially trade-disrupting wars. This simple framework allows us to derive several testable implications. First, RTAs are more beneficial to country pairs with a larger *probability* of war because the expected welfare gain of the political forum channel is higher. Second, recent *realizations* of war reduce the gain of an RTA because outbreaks of war increase the political costs of RTA negotiation. Third, trade gains from RTAs and the probability of war have a positive and complementary impact on RTA formation. The complementarity stems from the opportunity cost channel: the larger the trade gains, the larger the opportunity cost of a war and therefore the more useful an RTA is to secure peace which is more valuable to countries that have a higher probability of war.

Our empirical analysis is dedicated to estimating a model of RTA formation at the country-pair level over the 1950-2000 period to analyze whether the evolving geography of RTAs is consistent with the economic and political factors identified in the theoretical section. From the perspective of the identification strategy, a first concern is that many empirical determinants of wars and of the RTA-related trade gains are confounded: the gravity covariates, such as geographical distance, economic size, contiguity, cultural distance, etc., do affect the propensity to fight and the propensity to trade. This issue explains why the existing empirical literature on RTA formation has difficulty in disentangling the economic factors from the political factors. Here, we propose to rely on a theory-driven estimation procedure to quantify directly the potential trade gains generated by RTAs. To our knowledge our paper is the first to adopt such a strategy and this is an additional contribution of our paper. A second, and related, identification issue is that we need to differentiate between the probability of war, which increases the likelihood of an RTA, and the recent outbreaks of war, which reduces it. Our identifying assumption is that recent outbreaks are captured by the country-pair frequency of wars during the last 20 years, while probability can be measured by the country-pair frequency of old conflicts (over the period 1870-1945), a view which is consistent with existing evidence on the time-series autocorrelation of the war process.

Our empirical results, both in the cross-section and in the panel dimension, support our theoretical predictions. We find that trade gains and frequency of old wars have a high explanatory power and both increase the occurrence of RTA formation; their interaction term has also a positive impact and this confirms complementarity between economic and political factors. By contrast, recent war frequency decreases the occurrence of RTA formation suggesting the presence of windows of opportunity to lock-in RTAs. Finally we find that country pairs characterized by multilateral trade openness and a high

frequency of old wars are more likely to sign RTAs. We interpret this in the light of one our main findings in Martin, Mayer and Thoenig (2008) that multilateral trade openness, because it reduces bilateral economic dependence, increases the probability that a dispute escalates into a conflict. In other words, countries respond to the weakening of local economic ties (a side effect of globalization), and its peace-harming consequences, by reinforcing local economic ties through an RTA. From this point of view, we interpret RTAs as a logical political consequence of globalization.

We address the endogeneity issues by controlling for various codeterminants of political affinity, conflicts and trade; by including various country, country-pair, and year fixed-effects; and by instrumenting trade gains. All the results are robust to these different estimation strategies.

In the last section of the paper we quantify the identified mechanisms and perform several counterfactual experiments. We find that the complementarity between trade gains and the probability of war is sizeable and may even dominate the direct effect of each of this variable. This suggests that the opportunity cost channel is a first-order determinant of RTA formation. In other words, trade gains brought by RTAs are instrumentalized and are important as an intermediate objective of RTAs, their final goal being to pacify relations between countries. We also find that in a counterfactual world without any past history of warfare, the geography of RTAs formation would be radically different than the one actually observed. The same is true for a counterfactual world with no multilateral trade openness. We also check that our results are not driven by the European integration process although the mechanism is stronger for European country pairs.

The theoretical literature on RTA formation is very large. Nevertheless existing papers focus their analysis on the economic determinants only[5]; the role of security gains and military conflicts being largely ignored. From an empirical point of view, several papers study the economic determinants of RTAs (Baier and Bergstrand 2004, Egger and Larch 2008) under the identifying assumption that RTA-related trade gains are closely linked to the standard gravity covariates. As discussed above, this does not allow to discriminate between the economic and political factors, which is the purpose of our study. Symmetrically Mansfield and Pevehouse (2000) and Vicard (2009) look at the impact of RTA formation on the occurrence of military conflicts ignoring the potential role of economic factors.

The next section provides a simple theoretical framework and derives several testable implications. Section 3 presents the data, and discusses the empirical strategy. Section 4 reports the results and performs some quantification exercises, while section 5 concludes.

[5]This literature has analyzed the motives for building RTA mainly from a term-of-trade perspective (Bagwell and Staiger 1997, Ornelas 2005) and from a commitment perspective (Limao 2007, Maggi and Rodriguez-Clare 1998)

2 A simple framework

2.1 Timing and Welfare

We consider an insecure world where two countries decide whether to sign a bilateral RTA, which we interpret as a decrease in bilateral trade barriers with respect to the Most Favored Nation (MFN) tariff. We analyze hereafter how this decision is shaped by economic and political forces. To simplify the analysis, we focus, in this section, on two identical countries.

Two main characteristics describe the situation of bilateral relations between countries. First, whether they have signed an RTA or not. The variables of those who have signed an RTA are denoted with a superscript RTA; those who have not signed have no superscript. The second dimension is whether the two countries are at war or in peace.

The timing of events is as follows: in period 1, countries negotiate on the RTA. We make no particular assumption on the bargaining process but we assume that there is a political cost of negotiation C that is borne by each country. In period 2, we follow existing literature by assuming that a bilateral dispute may arise with probability δ for exogenous reasons (the existence of a common border, natural resources, ethnic minorities...) and may escalate into a military conflict with an endogenous conditional probability: e in absence of RTA or e^{RTA} if an RTA is in force. In period 3, economic gains are realized and each country gets an aggregate welfare level which depends on the existence of an RTA, and on the realization of a war at date 2.

In the rest of our analysis we express all welfare gains or losses as a percentage of a benchmark welfare, U_P, which is realized in the state of peace in absence of RTA. In this state both countries trade bilaterally and the MFN tariff level is applied. When war occurs, we assume that bilateral trade is fully disrupted and both countries go back to bilateral economic autarky. This trade disrupting effect of war is empirically well grounded (Blomberg and Hess 2006; Martin, Mayer and Thoenig 2008). Hence, welfare under war is given by $(1-W)U_P$ with $0 < W < 1$, whether an RTA is in force or not. The parameter W captures the direct costs of war (ie. destructions, death toll, etc.) augmented with the loss associated to bilateral economic autarky (with respect to the MFN situation). When a RTA is in force, additional welfare gains with respect to the MFN situation are generated only if peace is maintained; in that case welfare is given by $(1+T)U_P$. According to standard trade theory, $T > 0$ if the trade creation effect of the RTA dominates the trade diversion effect; otherwise $T < 0$. One of the purposes of the empirical analysis is to estimate precisely these trade gains associated to RTA formation.

The opportunity cost of war corresponds to the welfare differential between war and peace. From the previous discussion we see that in absence of an RTA, this differential is equal to WU_P while it is equal to $(W+T)U_P$ when a RTA is in force. As a consequence signing a RTA increases the

opportunity cost of a war by T/W percent.

2.2 Signing a RTA

At date 1, a RTA is signed when, for each country, the net expected surplus induced by the RTA is larger than its political cost. Noting V^{RTA} and V the *expected* welfare with and without RTA, the condition for RTA signature is:

$$V^{RTA} - V \geq C, \tag{1}$$

where $V = (1 - \delta e)U_P + \delta e(1 - W)U_P$ and $V^{RTA} = (1 - \delta e^{RTA})(1 + T)U_P + \delta e^{RTA}(1 - W)U_P$. Without loss of generality, we can express the political cost as a percentage of the benchmark welfare: $C = c \times U_P$ with $0 < c < 1$. Below, we detail some likely determinants of the negotiation cost c. Combining those equations with equation (1), the condition for signing an RTA becomes:

$$\underbrace{(1 - \delta e^{RTA})T}_{\text{trade gains}} + \underbrace{\delta \left(e - e^{RTA}\right) W}_{\text{security gains}} \geq c, \tag{2}$$

where on the LHS we have decomposed the net expected surplus of RTA into pure economic gains and security gains. Economic gains result from the increase in welfare from U_P to $(1+T)U_P$ when the RTA is active; however the RTA related trade gains T are realized only in periods of peace which occur with probability $(1 - \delta e^{RTA})$. The security gain of an RTA is associated with the potential decrease in the probability of escalation of disputes into war from e to e^{RTA} ; this allows to save on the costs of war W.

We now analyze the differential $(e - e^{RTA})$. As shown by the international relations literature (see Fearon 1995 and Powell 1999 for surveys), escalation to military conflicts can be interpreted as the failure of negotiations in a bargaining game. From this perspective, the probability of escalation depends negatively on the opportunity cost of war and positively on the degree of informational asymmetry between the two countries.[6] The rationale for the first channel is that, as the opportunity cost of war increases, countries have more incentive to make concessions in order to avoid the escalation of a dispute into a military conflict. The rationale for the second channel is that information asymmetries imply that during negotiations, countries do not report their true outside option, in order to extract larger concessions. This may prevent negotiations to succeed and disputes may escalate into war.

The signature of an RTA affects the probability of escalation, e, through these two distinct channels. First, as discussed in the previous section, RTA potentially increases the opportunity cost of war

[6] For a formal proof, see for example Martin, Mayer and Thoenig (2008) where we consider a fairly general bargaining game such that: (i) war is Pareto dominated by peace; (ii) countries have private information on the military and political strength of the other country; (iii) countries can choose any type of negotiation protocol. The negotiation is such that escalation to war is avoided whenever countries agree upon the sharing of the economic surplus under peace.

by T/W percent. This tends to reduce the probability of escalation. Second, regional integration produces a political spillover on conflict resolution by reducing the degree of informational asymmetries: Successful negotiations on economic and trade matters and the repeated interactions that follow these negotiations enable policy makers to learn about the other country. This channel has been discussed at length in the political science literature[7] , and many RTAs, such as the EU, ASEAN or MERCOSUR, have become venues to discuss political issues and potential disputes. Transposed to our model this discussion implies the following assumption on the probability of escalation under a RTA:

$$\frac{e^{RTA} - e}{e} = -\varepsilon_{pol} - \varepsilon_{cost}\frac{T}{W} < 0, \tag{3}$$

where ε_{pol} stands for the political spillover effect while $\varepsilon_{cost} > 0$ corresponds to the elasticity of escalation e to the cost of war. In the rest of the paper we refer to $(\varepsilon_{pol}, \varepsilon_{cost})$ as the security gains of RTA formation.

Under the reasonable assumption[8] that the RTA related trade gain T is small with respect to the cost of war W, we can combine (2) and (3) to get a first order Taylor approximation of the RTA signature condition (see appendix):

$$\Gamma \equiv T + \varepsilon_{pol}(\delta e \times W) + (\varepsilon_{cost} - 1)(\delta e \times T) \geq c, \tag{4}$$

where Γ corresponds to the benefits of RTA formation. This equation is our key theoretical relationship. It predicts that there are four determinants of signing RTAs:

- The first term on the left hand side of this inequality corresponds to the standard economic gains generated by the RTA on which the literature has focused. The larger these economic gains the higher the probability that the two countries sign it. The difficulty here is to produce a quantitative estimate of those trade gains for all country pairs. This is what we do in the empirical section.

- The second term corresponds to the positive political spillover of RTAs. Because signing an RTA allows to reduce the level of asymmetric information, it reduces the probability of escalation to war by ε_{pol} percent. Note that this political gain of RTAs is large when the potential welfare loss of war W is large and when the probability of war δe is also large.

[7]This argument, under the name of issue linkage, has been developed by political scientists working in the field of international relations, see Keohan and Nye (1977) , Haas (1980) and Mansfield and Pevehouse (2000).

[8]In the next section our empirical estimates show that the magnitude of T is approx. 1 percentage point of welfare. This is far below the existing estimates of the average cost of war W that can be found in the empirical literature (Glick and Taylor 2005)

- The third term is of ambiguous sign and depends whether the pacifying effect of RTAs through its impact on the economic opportunity cost of war is sufficiently large i.e. if $\varepsilon_{cost} > 1$. Two effects indeed go in opposite directions: on the one hand a high probability of conflict δe reduces the expected gain of an RTA because these gains are lost in times of war. On the other hand, a high probability of conflict also means that the pacifying effect of an RTA is very valuable. If policy makers believe that RTAs are indeed strong elements of pacification, this second effect should dominate. This is one of the questions that our empirical work will answer.

- The fourth term is the political cost of negotiation: it is linked to the current state of relations between the two countries and it can be affected by a myriad of institutional, political and historical factors.

2.3 Testable implications

From the previous discussion we are able to derive several testable implications. First consider a pure economic model where RTA formation brings no security gains. In our framework this corresponds to the case where the two elasticities $(\varepsilon_{pol}, \varepsilon_{cost})$ are equal to zero. And the RTA signature condition (4) becomes: $T - \delta e \times T - c \geq 0$. This condition yields our first testable implication:

Testable implication 1 *In absence of security gains, a larger probability of war, δe, reduces the probability of RTA formation*

The interpretation is straightforward. In absence of security gains, the only impact of a larger probability of conflict is to reduce the expected gain of an RTA because the RTA related trade gains are lost in times of war. This result offers a simple test of the prominent view held in the empirical literature, namely that RTA formation are determined by economic considerations only. If in our empirical analysis we find that the probability of war impacts negatively the probability of RTA signature, this is consistent with this view. Nevertheless our econometric results (see section 3.4) clearly pinpoint the opposite effect, namely a positive impact of the probability of war: this unambiguously rejects the "economics-only" view of RTA formation.

Turning back to the general case, the RTA signature condition (4) shows that the coefficient of the interaction term $(\delta e) \times T$ depends on the magnitude of ε_{cost}, the elasticity of the probability of escalation to the opportunity cost of war, with respect to 1. This leads to our second testable implication:

Testable implication 2 *Economic gains and the probability of war have a positive and complementary impact on the probability of RTA formation if and only if the costs of war deter sufficiently the process of escalation to war (i.e. $\varepsilon_{cost} > 1$).*

We assume that the political cost of negotiation, the parameter c in condition (4), depends on the number of recent wars. The reason is that recent military conflicts make it more difficult to engage in trade negotiations: grief brought by war generates vengeful feelings in the population which increase the political cost of such negotiations. Even though difficult to measure, vengeful feelings do exist and have been shown to die off gradually over time (see Mocan, 2008). This directly leads to our third testable implication:

Testable implication 3 *The frequency of recent wars reduces the probability of RTA formation*

The previous two results imply that the *probability* of war and the *realizations* of war potentially affect the probability of RTA formation in opposite directions. The probability of war, δe, tends to increase the expected gains of an RTA (when $(\varepsilon_{pol}, \varepsilon_{cost})$ are large enough) while the realizations of war increase the political cost of negotiation. From a time-series perspective, this means that the "overall" effect of war is ambiguous: pairs of countries which are prone to war do potentially benefit the most from RTA formation but are simultaneously those where the signature of such RTAs is politically difficult. A consequence for the empirical analysis is that our identification strategy must disentangle the impact of probability of war from the one of realization of wars.

We now consider the spillover effect of multilateral trade openness on the probability of RTA formation. One theoretical and empirical result in Martin et al. (2008) is that multilateral trade openness reduces the opportunity cost of a bilateral war and therefore makes it *more likely*. The rationale is that multilateral trade openness provides alternative trade partners and reduces bilateral trade dependence with the countries with which a dispute could escalate.

For a given level of trade gains T, the effect of multilateral openness can be interpreted in expression (4) as a decrease in the cost of war W. This itself increases the probability of escalation e. Hence, the net effect of multilateral openness on condition (4) is ambiguous. Straightforward computations show that the second effect dominates the first one if ε_{cost} is large enough (see the appendix for details):

Testable implication 4 *Multilateral trade openness and the probability of war have a positive and complementary impact on the probability of RTA formation if and only if the costs of war deter sufficiently the process of escalation to war (i.e.* $\varepsilon_{cost} > 1$ *).*

Political motives therefore imply that multilateral trade openness gives an incentive to sign RTAs to country pairs prone to conflict. The reason is that multilateral trade openness by reducing bilateral trade dependence reduces the cost of a bilateral conflict. An RTA is a way to compensate this potentially destabilizing consequence of multilateral trade openness and to bring back some bilateral trade dependence in the relation. This result supports the view that the development of multilateralism

during the 80s and early 90s could have triggered the wave of regionalism in the late 90s. This echoes a recent empirical finding by Fugazza and Robert-Nicoud (2009) of an emulator effect of multilateralism on regionalism for the US. They indeed find that the extent of post Uruguay Round PTAs (in term of included tariff lines) is positively affected by the extend of MFN tariff cuts negotiated by the US during the Urugay Round. While Fugazza and Robert-Nicoud provide no theory for their intriguing finding, our results suggest that this emulator effect of multilateralism on regionalism could be driven by security purposes. Finally note that results 2 and 4 allow to test internal consistency of our framework. Indeed both depend on the sign of $(\varepsilon_{cost} - 1)$. In our empirical section, if we find some evidence for complementarity between economic gains and war, we should find evidence for complementarity between multilateral trade openness and war as well. The opposite finding would refute our theory.

3 Empirical Analysis

3.1 Data

There are two main parts to the empirical investigations of this paper. We first estimate the economic gains of RTA, which involves essentially running a gravity equation over a sufficiently long time period to be able to identify the trade creation in the within dimension. We make use of the gravity dataset constructed for Martin et al. (2008) and Head et al. (2008), and described in greater detail in those two papers. Essentially, any gravity dataset requires source data for a trade flow variable, and a list of gravity controls. The trade flow source is IMF DOTS, with a procedure to extract the most possible information from mirror flow declarations.

The list of gravity controls includes the classical bilateral distances, contiguity, colonial linkages, and a common (official) language dummies. All those come from the CEPII distance database (`http://www.cepii.fr/anglaisgraph/bdd/distances.htm`). Later in the paper we also use a common legal origin dummy available from Andrei Shleifer at `http://post.economics.harvard.edu/faculty/shleifer/Data/qgov_web.xls`, and a variable for bilateral genetic distance, available from Spolaore and Wacziarg (2009).

More central in our case are the regional trading agreements. RTAs are constructed from three main sources: Table 3 of Baier and Bergstrand (2007) supplemented with the WTO web site (`http://www.wto.org/english/tratop_e/region_e/summary_e.xls`) and qualitative information contained in Frankel (1997).

The RTA dummy is also the dependent variable of our second and main empirical exercice, which explains their formation. In those regressions, our main variables of interest are related to old and new wars. The source data for military conflicts is the Correlates of War project (`http:`

//www.correlatesofwar.org/). More precisely, we use the information contained in the Militarized Interstate Disputes database that lists all bilateral interstate conflicts from 1816 to 2001, and quantifies their intensity on a 1 to 5 scale[9] (for a precise description of the source data, see Martin et al. 2008). We concentrate on the 1870-2001 period because 1870 is essentially the time when most European countries have a stabilized geographical and political structure. The old wars variable calculates the percentage of years with active military conflicts between the two countries, during the 1870-1944 period. This creates an immediate problem with countries that did not exist in this period naturally. What is the historical war propensity of the pair Algeria - Nigeria for instance? Absent detailed information impossible to gather for all pairs of ex-colonies and all years prior to independence, we envision several strategies, which range from assuming peace to dropping those observations. Those strategies and results are detailed below in the results section. Recent wars are taken to be the the same percentage of military conflicts, but for a moving window of 20 years before the year under consideration. For both variables, we consider only the two most severe types of wars, coded 4 and 5 in the COW database (see Martin et al. 2008 for examples).

In those regressions, there are other bilateral political variables, which serve as controls in the list of RTA determinants. Those include the correlation of roll-call votes recorded for the two countries in the General Assembly of the United Nations (from Gartzke et al., 1999), a dummy for the existence of a military alliance (from COW), and the sum of democracy indices (from Polity IV)

3.2 Estimating the Economic Gains of RTA

The objective of our empirical analysis is to estimate equation (4). The first task is to obtain $\hat{T}$, the estimate for the economic gains of a RTA. The existing literature on RTA formation (Baier and Bergstrand 2004, Egger and Larch 2008) proxies those gains with the standard gravity covariates, such as economic size, geographical distance, remoteness, contiguity, etc. in a reduced-form estimation of RTA formation. Given that our purpose is to understand the relation between economic and political factors, we cannot follow the same route. Indeed it is extremely likely that the gravity covariates affect both economic and political factors. Hence we must rely on a theory-driven empirical strategy to assess the economic gains of RTA formation and to disentangle them from the political factors.

Let us consider the wide class of trade models where aggregate welfare is derived from a CES utility function.[10] We now use subscripts for countries and time.[11] Country i welfare at date t is given

[9]The scale is the following: 1 = No militarized action, 2 = Threat to use force, 3 = Display of force, 4 = Use of force, and 5 = War, defined as a conflict with at least 1000 deaths of military personnel.

[10]Dixit-Stiglitz-Krugman (DSK) monopolistic competition approach is an example of such modelling, the national product differentiation approach of Anderson and van Wincoop (2003) is another example.

[11]One adjustment we must make to our theory when applying it empirically, is to account for heterogeneity between different country pairs, and variance in the time dimension.

by

$$U_{it} = E_{it}/P_{it}, \tag{5}$$

where E_{it} is nominal GDP and P_{it} is the price index. The price index can be written as

$$P_{it} = \left[\sum_{k} \mu_{kt} \tau_{kit}^{1-\sigma}\right]^{1/(1-\sigma)}, \tag{6}$$

where σ is the elasticity of substitution between goods, μ_{kt} stands for all factors in the model that makes country k a good exporter[12] and $\tau_{kit}^{1-\sigma}$ represents bilateral trade freeness, where $\tau_{kit} > 1$ is the iceberg-type price shifter which accounts for all trade barriers. In this context, bilateral trade obeys the following gravity equation governing imports of i from j in year t:

$$m_{jit} = \mu_{jt} E_{it} P_{it}^{\sigma-1} \tau_{jit}^{1-\sigma} \tag{7}$$

We seek to understand the welfare impact of a RTA formation between countries i and j in a partial equilibrium framework. The general equilibrium case is a fascinating issue but raises complexities that go far beyond the scope of this paper. In particular, general equilibrium has to take into account the relocation effects of all firms in the world following each signing of an RTA. Potentially, the wages in all countries can be affected as well, both effect affecting the whole distribution of nominal GDPs. Moreover the drop in tariff revenues following RTA formation affects negatively aggregate income. The economic geography literature has shown that considering those effects requires numerical simulations, since no analytical solution emerges in a multiple country world of that complexity. We stick to a framework that maintains analytical solutions that can be brought to the data. Our identifying assumption is therefore that in equation (5) the price index P_{it} is affected by the RTA while E_{it} remains unchanged. Hence our procedure considers only the price effect of RTA formation. However we see no reason for the induced measurement errors on the estimated welfare gains to be correlated with our covariates of interest, detailed in equation (15). Hence this restrictive assumption is unlikely to contaminate the rest of our econometric analysis.

The level of P_{it} depends on the existence of a RTA through the bilateral trade barriers in equation (6):

$$\tau_{jit} \equiv \exp(-\rho \mathrm{RTA}_{jit}) \eta_{jit}, \tag{8}$$

where η_{jit} is the residual component of trade costs while RTA_{ijt} is a dummy variable set equal to 1 when a RTA is in force between i and j in t. The parameter ρ depends directly on the preferential tariff cut.

[12] In the DSK model for instance, this term is $n_k p_k^{1-\sigma}$, a positive function of the number of varieties, and negative one of the price charged by firms located in k.

We can combine equations (5) and (6) to obtain T_{ijt}, the percentage change in utility of i following a RTA with j :

$$T_{ijt} = \left[\frac{\sum_k \mu_{kt} \eta_{kit}^{1-\sigma}}{\mu_{jt} \exp[(\sigma-1)\rho] \eta_{jit}^{1-\sigma} + \sum_{k \neq j} \mu_{kt} \eta_{kit}^{1-\sigma}} \right]^{1/(1-\sigma)} - 1 \tag{9}$$

We actually estimate this equation using bilateral trade data over the 1950-2000 period (see section 3.1 for the data description). This requires several steps. First, we can use our definition of trade costs (8) in the gravity equation (7) to obtain a new version of the gravity equation

$$\ln m_{jit} = \ln \mu_{jt} + \ln \left(E_{it} P_{it}^{\sigma-1} \right) + (\sigma-1)\rho \mathrm{RTA}_{jit} + (1-\sigma) \ln \eta_{jit} \tag{10}$$

This leads to estimating a panel specification of the gravity equation (10):

$$\ln m_{jit} = FX_{jt} + FM_{it} + \beta \mathrm{RTA}_{jit} + u_{jit} \tag{11}$$

where u_{jit} is the error term, FX_{jt} is an (exporter×year) fixed effect, and FM_{it} is an (importer×year) fixed effect. This specification has the advantage of remaining flexible in terms of the exact underlying trade model, while enabling to extract the parameters of interest for the calculation of the utility change in (9).[13] Indeed, comparing (10) and (11), one obtains $\mu_{jt} = \exp(\widehat{FX}_{jt})$, $\exp((\sigma-1)\rho) = \exp(\hat{\beta})$, and $\eta_{jit}^{1-\sigma} = \exp(\hat{u}_{jit})$. Our point estimate of $\hat{\beta}$ is 0.311, yielding a predicted increase in bilateral trade of 37% from entry into a RTA. For comparison purposes, Baier and Bergstrand (2007) using bilateral fixed effects and year dummies on a panel (for every five years) from 1960–2000 find an estimate of 0.68 (last column of Table 4). Head et al. (2009) find 0.378 using their tetradic method which is most comparable with the method used here.

Our second step retrieves those point estimates and substitute them into equation (9). This gives us our empirical estimator of the economic gains of RTA:

$$\hat{T}_{ijt} = \left[\frac{\sum_k \exp(\widehat{FX}_{kt} + \hat{u}_{kit})}{\exp(\hat{\beta} + \widehat{FX}_{jt} + \hat{u}_{jit}) + \sum_{k \neq j} \exp(\widehat{FX}_{kt} + \hat{u}_{kit})} \right]^{1/(1-\sigma)} - 1, \tag{12}$$

where we use the standard calibration for the elasticity of substitution in the empirical trade literature $\sigma = 5$.[14]

Figures 1 and 2 and Table 1 describe our trade gains variable $\hat{T}_{ijt}$.

[13] Our panel contains bilateral trade flows over the 1950-2000 period. We exploit the within dimension of this dataset, in order to identify $\hat{\beta}$ from entries and exits into the agreements rather than from a comparison across country pairs. Thus, in (11), we allow u_{jit} to be additively decomposed into a time-invariant and a time-varying element. The regression also includes year dummies.

[14] GTAP version 5, the workhorse model for computable general equilibrium analysis of trade liberalization retains an average estimate of 5.3 (Dimaranan and McDougall, 2002). Econometric evidence by Hertel et al. (2007) point to an average elasticity of substitution of 7.0, while Broda and Weinstein (2006) estimate a mean σ of 4.0 for their most recent period and a 3-digit classification (their Table IV).

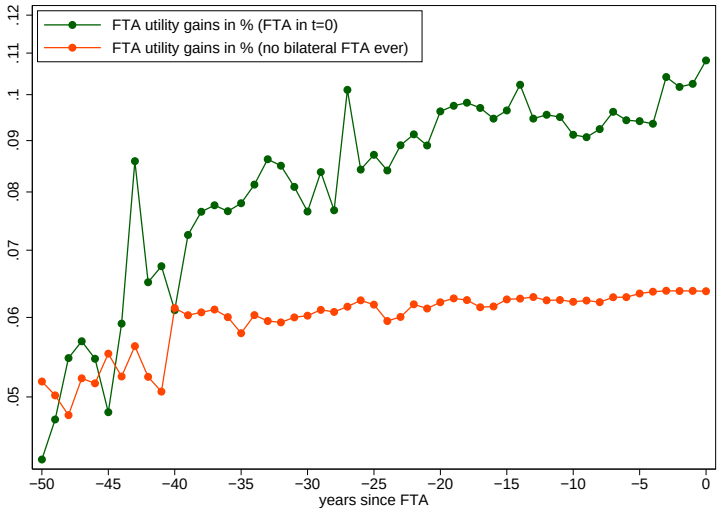


Figure 1: Utility gains RTA / no bilateral RTA

In figure 1, we plot the average estimated trade gains of joining a RTA for two types of country pairs: i) those that do enter a bilateral RTA at some point in our sample, ii) those that do not. For the second group, we want to make it as comparable as possible to the first one, and therefore, we keep only those country pairs where both members do enter an RTA with a third country but do not sign a bilateral one.[15] On the horizontal axis, we show the number of years before the signature of the bilateral RTA for those who sign it and the number of years until year 2000 for the control group. The main result is the difference in trends: the RTA signatories have estimated trade gains that grow as we get closer to the actual signing, whereas nothing visible happens in the control group. This suggests that our measure of economic gains from a RTA can be used as a predictor of the decision to enter a bilateral RTA, both in the cross-section in the years before the signature, and in the within dimension, looking at when countries decide to sign.

Figure 2 focuses on the set of countries that do enter a RTA, and distinguishes the European Union members (defined as EU15) from others. We are also able to look at what happens to our measure of trade gains after the RTA signature. One can observe that the trend before signature continues afterwards. This is not surprising: RTA gains come from trade creation, and it is therefore logical that comparing the utility before and after the RTA implementation reflects the amount of trade created within the pair. Hence there is potentially a reverse causality from RTA formation on the trade gains. This points to an important methodological issue that we have to address in our econometric analysis.

In table 1 we report the estimated trade gains in 1956, one year before the Rome Treaty, for the

[15]This restriction does not affect radically the shape of the curve. When comparing with the whole set of country pairs which do not sign a bilateral RTA, the graph looks almost the same.

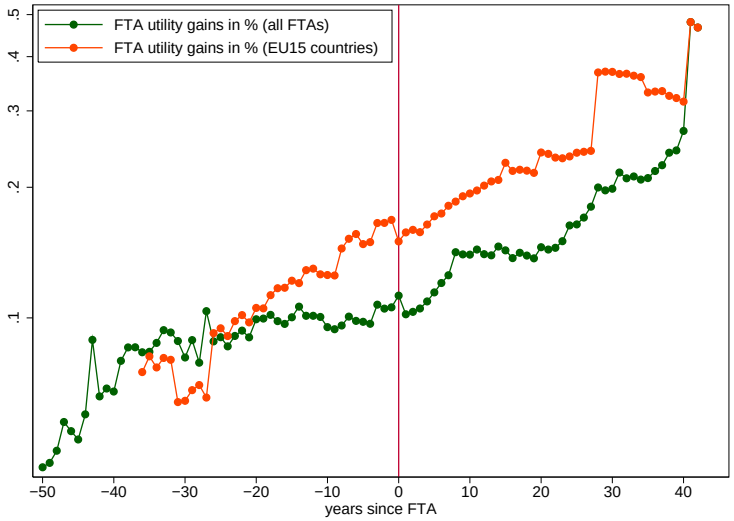


Figure 2: Utility gains average RTA / EU15

subsample of 50 country-pairs (out of a sample of 8240) for which the trade gains are the largest. We report the country-pair minimum, $min(\hat{T}_{ijt}, \hat{T}_{jit})$ and the country-pair unweighted average, $(\hat{T}_{ijt} + \hat{T}_{jit})/2$. There may be a large discrepancy between these two figures, especially in asymmetric country-pairs where the smallest country tends to gain much more than the biggest country. In our econometric specifications we focus on the country-pair minimum because it is the theoretically grounded one when welfare transfers and compensation schemes between trade partners are difficult to implement after RTA formation. The interpretation of the table is the following: in 1956, USA and Canada would have increased their welfare at least by 2.2 percent if they had formed a RTA. [16].

A look at the table and figures of this section confirms that the estimated trade gain are small. For instance, in figure 1, our estimate of the average gain from entering a RTA (at the year of signature) is 0.13%. This order of magnitude is consistent with standard results of trade gains estimates based on Computable General Equilibrium (CGE) analysis. A recent example evaluating the impact of the Free Trade Agreement of the Americas by Hertel et al. (2007) finds an estimate of average utility changes for potential members at 0.25% (their Table 5). Recall that our framework omits the effects that RTAs have on nominal GDP, possibly explaining the smaller effects than those found in full-fledged general equilibrium modeling.

Another striking feature of our results is that, one year before the Rome Treaty, the country-pairs composed of the EEC founding countries (in bold) are in the group of large trade winners, but not systematically among the top ones. This is in line with our view that economic gains are important

[16]Regarding this USA-Canada example, the percentage increase in welfare is 2.2% for the USA and 6.1% for Canada such as the country-pair average increase is 4.15%

determinants of RTA formation; and the last column of the table reinforces that impression since many of those pairs will end up having a RTA in the end. However it also suggests that those economic gains are not the only determinants of RTA formation. Political issues do matter. We also add to the table geographical weighted distance between those two countries, and the minimum of their bilateral import / GDP ratios, both being in fact closely related to utility gains in CES-type models of trade.

3.3 Empirical Model and Identification Strategy

In this section we present our strategy to identify the causal links from economic and security gains from RTA formation. Using our theoretical equation (4), we consider a specification where, for a country-pair (i, j) at year t, a regional trade agreement is signed when:

$$\Gamma_{ijt} > \epsilon_{ijt}. \tag{13}$$

In this equation, Γ_{ijt} is the utility gain from signing the agreement, and ϵ_{ijt} corresponds to the negotiation cost, c in our theory. Empirically, ϵ_{ijt} is the unobserved component of the decision process, submitted to stochastic shocks in political affinity for instance, which transforms (13) into a probability of RTA formation. The functional form taken by this probability depends upon the distribution assumed on ϵ_{ijt}. With a Gumbel /Type I extreme value distribution (see chapter 3 of Train, 2003 for an example of the logit formula derivation), we obtain the logit probability to be estimated using maximum likelihood:

$$\mathbb{P}(\text{RTA}_{ijt} = 1) = \frac{\exp(\Gamma_{ijt})}{\exp(\Gamma_{ijt}) + 1}. \tag{14}$$

The dependent variable RTA_{ijt} is a dummy coding for the existence of a RTA between i and j in year t, and Γ_{ijt} follows from equation (4):

$$\Gamma_{ijt} = \alpha + \beta_1 \min(\hat{T}_{ijt}, \hat{T}_{jit}) + \beta_2 WAR_{ij} + \beta_3 \min(\hat{T}_{ijt}, \hat{T}_{jit}) \times WAR_{ij} + \mathbf{Z}_{ijt}\boldsymbol{\beta}. \tag{15}$$

The RTA-induced economic gains are measured by $\min(\hat{T}_{ijt}, \hat{T}_{jit})$, the country-pair minimum of gains; this is a natural consequence of our theoretical setup where RTA must be Pareto-improving in absence of any compensatory transfers within the country-pair.[17] In our robustness analysis we test for the possibility of additive transfers by measuring trade gains with the country-pair average $(\hat{T}_{ijt}+\hat{T}_{jit})/2$ rather than the minimum. WAR_{ij} is a proxy for the probability of war (see the discussion below on how we measure this variable), $\mathbf{Z}_{ijt}$ is a set of control variables and the residual term ϵ_{ijt} captures cross-dyadic heterogeneity in the political costs of negotiation.

[17] In our theoretical setup the two countries i and j are assumed to be symmetric for the sake of exposition. Relaxing this assumption and ignoring compensatory transfers, the condition (4) is now country-specific given that the trade gains (T_{ij}, T_{ji}) are potentially asymmetric. A RTA is formed when the minimum of the two country-specific conditions (4) is positive.

Table 1: Estimated Trade Gains for the top 50 country-pairs in 1956

Country pair		Trade gains		bil. open.	dist.	ever fta?
		min T	mean T	min $\frac{\text{imports}}{\text{GDP}}$	kms	
SUN	CHN	1.95%	2.919%	.622%	5507	No
USA	CAN	1.786%	3.399%	.748%	2079	Yes
NLD	**BEL**	**1.054%**	**1.261%**	**4.38%**	**161**	**Yes**
CZS	SUN	1.031%	1.891%	.323%	2388	No
POL	SUN	.741%	1.715%	.231%	2067	No
SYR	LBN	.667%	1.064%	2.917%	228	No
CAN	GBR	.637%	.718%	1.661%	5850	No
ROM	SUN	.617%	2.294%	.192%	2142	No
FRA	**DEU**	**.57%**	**.789%**	**1.019%**	**790**	**Yes**
POL	CZS	.568%	.701%	.743%	387	No
NLD	**DEU**	**.564%**	**.976%**	**1.009%**	**379**	**Yes**
GBR	AUS	.546%	1.899%	1.128%	16602	No
BEL	**FRA**	**.546%**	**.754%**	**.559%**	**526**	**Yes**
BRA	ARG	.498%	.555%	.855%	2392	Yes
USA	GBR	.488%	.713%	.199%	6878	No
USA	BRA	.469%	1.346%	.191%	8089	No
GBR	NZL	.457%	2.165%	.942%	18521	No
USA	VEN	.444%	2.249%	.181%	4204	No
FRA	MAR	.424%	1.986%	.433%	1706	Yes
SUN	FIN	.385%	.665%	.119%	1635	No
BGR	SUN	.381%	1.84%	.118%	2391	No
BEL	**DEU**	**.38%**	**.789%**	**.677%**	**423**	**Yes**
FRA	IRQ	.376%	.384%	.383%	3805	No
CZS	CHN	.369%	.429%	.161%	7790	No
DEU	SWE	.361%	1.017%	.643%	929	Yes
USA	JPN	.352%	1.49%	.143%	10286	No
DEU	**ITA**	**.346%**	**.671%**	**.615%**	**1014**	**Yes**
AUT	ITA	.338%	.479%	.506%	701	Yes
GBR	SWE	.337%	.702%	.692%	1293	Yes
GBR	IND	.329%	1.161%	.676%	7324	No
GBR	NLD	.319%	.483%	.657%	468	Yes
HUN	SUN	.319%	1.066%	.098%	2334	No
USA	DEU	.312%	.713%	.127%	7595	No
JPN	PHL	.301%	.535%	.432%	2957	No
SWE	NOR	.29%	.676%	.766%	503	Yes
USA	CUB	.289%	2.737%	.118%	2581	No
POL	CHN	.287%	.288%	.125%	7457	No
GBR	DNK	.285%	1.008%	.585%	920	Yes
IRN	IND	.274%	.362%	.235%	2916	No
NLD	**FRA**	**.274%**	**.276%**	**.284%**	**661**	**Yes**
SAU	JPN	.273%	.315%	.512%	8854	No
ITA	SAU	.273%	.323%	.408%	3586	No
CHE	DEU	.273%	1.024%	.484%	543	Yes
JPN	IND	.267%	.349%	.372%	6003	No
SWE	DNK	.266%	.464%	.703%	450	Yes
USA	MEX	.264%	2.733%	.107%	2468	Yes
NLD	SWE	.261%	.402%	1.433%	1009	Yes
GBR	FRA	.261%	.337%	.422%	750	Yes
NOR	DNK	.26%	.263%	1.047%	560	Yes
CHE	ITA	.26%	.485%	.388%	610	Yes

Note: Lines in boldface indicate pairs that sign the Rome Treaty establishing the European Economic Community a year later.

In equation (15) we expect β_1 to be positive. The coefficient β_2 tests for the existence of a political spillover of RTA. It is expected to be nonnegative. The interpretation of the sign of β_3, the coefficient of the interaction term, can be misleading in a logit specification due to the non-linearity of this model (see Ai and Norton 2003). The logit specification also makes the handling of panel data techniques such as within estimation more complicated, while the marginal effects tend to be similar to the Linear Probability Model (LPM) in many cases as stated in Angrist and Pischke (2009, p107). Hence in all specifications of (15) where the interaction term is included, we estimate a linear probability model rather than a logit model. This standard choice also facilitates the interpretation of the coefficient.[18] In that case the coefficient β_3 corresponds to a marginal effect and it can be simply interpreted as a test of (complementarity/substitutability) between economic and security gains: from result 2, complementarity ($\beta_3 > 0$), is expected when the opportunity cost channel is at work (i.e. the pacifying effect of RTAs is large so that $\varepsilon_{cost} > 1$). Interestingly, we can infer from these estimates the underlying value of ε_{cost} that comes from the observed geography of RTAs. Comparing equations (4) and (15), we obtain $\hat{\varepsilon}_{cost} = 1 + \hat{\beta}_3/\hat{\beta}_1$. This estimate can give us an indication on the view of policy makers on the pacifying effect of RTAs as revealed by the choice they made on RTA formation.

3.3.1 Measuring conflictuality

In equation (15), there are two central variables, $\hat{T}$ and WAR. The construction of $\hat{T}$ is detailed in section 3.2, we know turn to our measurement of war probability, WAR. A natural proxy for this probability is the historical frequency of wars between each country pair. However, there are issues with this way of measuring WAR.

The discussion following result 3 states that although the *probability* of war tends to make RTA formation more likely, the *realization* of war, by increasing the political cost of negotiation, tends to make RTA formation less likely. Therefore if we measured WAR with the country-pair historical frequency of wars, the two channels would be mixed and the estimated coefficient would capture the net effect of the two mechanisms. The sign of this net effect could then be either positive or negative.

The identifying assumption we use to separate the impact of underlying war probability from the one of actual conflicts (realizations), is that war realizations raise the cost of subsequent bilateral negotiations, and that this cost exhibits a decay over time. By contrast, as supported by empirical evidence, we assume that bilateral war probability is more stable over time. One way to think about this is that the feelings of revenge and grievance that follow a war are most vivid just after a war and then "depreciate" over time. This assumption can be justified by the recent empirical findings of

[18] However, an area where logit (or probit) is undoubtedly preferable to LPM is the predictions one can make when changing one or more variables more than marginally. Probabilities have to be bounded between 0 and 1 by the model then in order to yield meaningful predictions. In our quantification exercise, we hence return to the logit specification.

Mocan (2008) who shows that, for victimized individuals, the desire for revenge and the propensity to punish decreases through time. On the other hand, a very robust finding of the empirical literature on conflicts is that the frequency of old wars is a strong predictor of the frequency of current wars (Collier et al. 2004): This result stems from the existence of important time-invariant determinants of disputes and war that can be observed or not by the econometrician (common border, natural resources, ethnic rivalries, cultural antagonism etc.).

In light of this discussion we implement the following identifying strategy: We proxy the probability of war at date t, WAR_{ijt}, with the country-pair frequency of bilateral wars which occurred between 1870 and 1945. We call it frequency of *old wars*. This proxy being time-invariant, we suppress the time index, which gives the variable WAR_{ij} in the econometric equation (15). In addition, we proxy the realization of wars with the country-pair frequency of bilateral wars which occurred during the last 20 years: We call it frequency of *recent wars*. If this strategy is relevant we should observe the marginal effect of recent wars to be negative as it captures the political cost of realized conflicts. The marginal effect of old wars should be positive because it mainly captures the probability channel through which RTAs are more beneficial to country pairs with a high propensity to conflicts. In our robustness analysis we test definitions of old and recent wars with alternative time spans.

3.3.2 Endogeneity issues

The estimates of our main coefficients of interest, $\beta_1, \beta_2, \beta_3$, are potentially contaminated by several sources of endogeneity. For each of them, we now explain how we deal with it.

Measurement errors

As seen in the preceding subsection, our approach to measure WAR_{ij} has the advantage of purging the effect of recent realization from the impact of war probability that we intend to capture. However, by relying on old history of conflicts, it introduces noise in the measurement of current war probability. Indeed, some causes of disputes in the late 19th century (e.g. the building of colonial empires) have lost their explanatory power. Simultaneously, new causes have emerged in the late 20th century (oil production; water supply access; religious tensions). Those time-varying determinants imply measurement error in the current probability of war: This should go against our results by inducing a bias towards zero in the estimated coefficients of interest.

As discussed in section (3.2), our estimate of the trade gains, $\hat{T}_{ijt}$, relies on a partial equilibrium analysis and are therefore a noisy measure of T_{ijt} the true trade gains. This could bias β_1 and β_3, the two coefficients involving $\hat{T}_{ijt}$, towards zero, underestimating the magnitude of the effects we are trying to identify. However we see no particular reason why the measurement error $(\hat{T}_{ijt} - T_{ijt})$, which

is part of the error term, should be correlated with our other variable of interest, WAR_{ij}, which should leave the estimate of β_2 unaffected by this measurement error problem.

Reverse causality

Figure 2, highlights the possible reverse causality link from RTA to trade gains *following* RTA formation. In order to eliminate this issue that can overestimate the coefficient β_1, we need to compare $\hat{T}_{ijt}$ across country pairs or time *before* the agreement actually takes place. Similarly, this reverse causality issue may bias downwards β_2 because RTA formation is likely to reduce the probability of future conflicts.

In the cross-section dimension we thus estimate equation (15) in year $t = 2000$ for dyads where a RTA does not exist in 2000. For dyads where the two countries are members of a RTA in 2000, their RHS variables are set to their values one year before the RTA formation. For example, in the case of USA-Canada, this means that all the RHS variables take their 1988 values. This methodology generalizes the approach by Baier and Bergstrand (2004) and allows to control for reverse causation. Correspondingly, in the panel estimates of (15), we focus on "RTA onset", that is we analyze, for each dyad, years up to the signature of the RTA, dropping observations after the signature. This is very similar to the method used by researchers studying the determinants of conflicts (Fearon 2005).

Omitted variables

In equation (15), the coefficients of economic gain and of its interaction term with war, β_1 and β_3, could be contaminated by omitted co-determinants of economic gains, $\hat{T}_{ijt}$, and of unobserved political costs of RTA formation, ϵ_{ijt} (i.e. the residual). This may arise because the structural relationship (12) defining $\hat{T}_{ijt}$ depends on $\hat{u}_{jit}$, the estimate of (logged) bilateral trade freeness retrieved from the auxiliary gravity equation (10). Indeed, several determinants of bilateral trade freeness (or conversely trade barriers) might also affect the bilateral political affinity and consequently the political costs of RTA formation (e.g. commonality of language and culture, economic embargo, etc.). A striking illustration is provided in Michaels and Zhi (2007) who show that the deterioration of political relations between the US and France over the 2002-2006 period resulted in a significant increase in their bilateral trade barriers following changes in attitudes towards France in the US.

To address this concern, we first add to the set of control variables $\mathbf{Z}_{ijt}$ a series of co-determinants of bilateral trade barriers and political costs of negotiation. This encompasses the standard time invariant gravity controls (distance, contiguity, common language, etc.) and various time-varying proxies of bilateral political affinity such as a dummy variable coding for the existence of a military alliance, a measure of bilateral correlation in UN votes from Gartzke et al. (1999) and lastly the

country-pair sum of democracy indices from the Polity IV database. Indeed, the democratic peace hypothesis, which has been studied by both political scientists and economists (see Levy and Razin, 2004, for a recent explanation of the hypothesis) states that democratic countries are less prone to violence. But democratic countries are also more open to trade. In the panel specifications, we can be more general in those controls, by including a country-pair fixed effect to purge from remaining time-invariant unobserved heterogeneity.

In spite of all these controls, we cannot rule out the possibility that the coefficient of trade gains, β_1, is still contaminated by unobserved *time-varying* co-determinants of bilateral trade freeness, $\hat{u}_{jit}$, and political affinity, ϵ_{ijt}. To solve this last problem, we directly include $\hat{u}_{jit}$ as a control variable. This strategy allows to identify β_1 by exploiting the variations in trade gains $\hat{T}_{ijt}$ net of $\hat{u}_{jit}$. This solves the omitted variable problem because those variations are not driven by bilateral shocks and so cannot be correlated with the (residual and unobserved) political costs of negotiations ϵ_{jit}. Indeed a look at the structural relationship (12) makes it clear that those variations are driven by changes in the exporter fixed effects $\widehat{FX}_{kt}$. This strategy is in fact akin to an control function approach where the trade gains $\hat{T}_{ijt}$ are instrumented with a remoteness index based on the exporter fixed effects $\widehat{FX}_{kt}$ (see Imbens and Wooldridge, 2007).

Regarding β_2 and β_3, the coefficients of the probability of war and of its interaction term with trade gains in the econometric specification (15), the omitted variable problem is potentially severe: Indeed any time-invariant determinant of the unobserved political costs of RTA formation ϵ_{ijt}, is also likely to affect the underlying probability of war, WAR_{ij}. For example, disputes linked to common borders, natural resources, migration waves, etc., are likely to increase the underlying probability of war and make negotiation on RTA formation politically more costly. This suggests that the omitted variable problem should induce a downward bias which goes against our hypothesis. Note that the various gravity and political affinity controls included in $\mathbf{Z}_{ijt}$ are likely to absorb most of the cross-sectional variations in bilateral disputes. Moreover, we also include as a control variable a measure of bilateral genetic distance. Indeed recent findings by Spolaore and Wacziarg (2009) show that genetic relatedness has a positive effect on bilateral conflict propensities in the cross-section. This is because more closely related populations, on average, tend to interact more and develop more disputes over sets of common issues. Hence we expect genetic distance to reduce the probability of war and to increase the probability of RTA formation. Finally, in our panel estimates, we include country-pair fixed effects. This makes impossible the identification of β_2, the coefficient of the time-invariant variable WAR_{ij}. Nevertheless, we can still estimate β_3 which is now immune to the omitted variable bias: There is indeed no particular reason for the determinants of political costs ϵ_{ijt} to have a larger effect on RTA formation in dyads where the trade gains are larger.

4 Results

4.1 Econometric estimates

We start in Table 2 with a cross-sectional analysis of RTA determinants. By cross-sectional we mean that we take the world in the year 2000, and attempt to explain which of the country pairs are in a RTA. Some determinants will be time invariant (e.g. distance), some will have a time dimension. For the latter set of variables, we consider the variable for the year immediately preceding the signature of the RTA. For instance trade gains are taken in 1956 (the year before the Rome Treaty) for the Franco-German case, and in 1993 (the year before NAFTA) for the USA-Mexico one. Since this variable is calculated as a percentage of utility, we have a variable that is relevant at the moment of the decision, while maintaining comparability across observations.

Our first column runs logit on a very simple specification which has only the log of the estimated trade gains [19] and the frequency of old wars as covariates. As expected both enter positively, with a quite large overall explanatory power, and a high degree of statistical significance.[20] The fact that our two main variables of interest are sufficient to explain more than a quarter of the observed variance in RTA formation seems to provide encouraging empirical support to our theory-driven construct of trade gains. In this first column, the old war variable WAR_{ij} is restricted to the small number of dyads which exist before 1945. In particular, all country pairs that involves a former colony (India-Japan, Germany-Ivory Coast for instance) are dropped from this regression. In column (2) we adopt the following alternative strategy: We set WAR_{ij}, the old war variable, to 0 for country-pairs which did not exist before 1945; we also include a dummy variable coding for those pairs. As can be seen from the comparison of columns (1) and (2), the two variables of interest have very close coefficients with this procedure and the fit is very comparable, which makes us confident that it does not alter our results while augmenting substantially the number of observations[21] . We maintain this procedure throughout.

Column (3) introduces bilateral trade freeness. As stated above, this is intended to circumvent any contamination of the coefficient on trade gains, by unobserved co-determinants of bilateral trade freeness and political affinity. We therefore control for $\hat{u}_{jit}$, the estimate of bilateral trade freeness obtained from the gravity equation (10). As expected, this variable enters positively and results in a decrease of the effect of trade gains as it purges from contemporaneous bilateral affinity which causes both the probability of signing a RTA and the trade gains to be high.

[19] We take the log because of the left-skewness of the distribution of estimated trade gains

[20] Interpreting the economic magnitude of coefficients is here slightly more complex than in a simple linear setting. We devote the next section to it.

[21] It can be noted that those non-existing dyads, mostly combinations of colonies at the end of WWII, have been less involved in the RTA movement, as revealed by the negative coefficient of the dummy variable.

Table 2: RTA determinants, benchmark regressions

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var.	FTA	FTA	FTA	FTA	FTA	FTA	FTA	FTA	FTA onset
Period	2000	2000	2000	2000	2000	2000	2000	1950-2000	1950-2000
trade gains ($\hat{T}_{ijt}$)	0.553^a	0.415^a	0.330^a	0.016^a	0.014^a	0.013^a	0.009^a	0.007^a	0.002^a
	(0.038)	(0.019)	(0.021)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
war freq. pre-1945 (WAR_{ij})	7.840^a	8.328^a	9.257^a	1.912^a	5.169^a	3.823^a	7.963^a		
	(1.337)	(1.271)	(1.273)	(0.116)	(0.341)	(0.390)	(0.757)		
dyad did not exist pre-1945		-0.783^a	-0.902^a	-0.053^a	-0.063^a	-0.034^a	-0.045^a		
		(0.107)	(0.109)	(0.007)	(0.007)	(0.008)	(0.008)		
ln bil. trade freeness			0.217^a	0.010^a	0.010^a	-0.009^a	-0.011^a	-0.003^a	-0.002^a
			(0.025)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.000)
trade gains × wars pre-1945					0.354^a	0.309^a	0.460^a	0.163^a	0.062^a
					(0.035)	(0.038)	(0.042)	(0.026)	(0.011)
war freq. $[t-20; t-1]$						-0.262^a	-0.441^a	-0.067^a	-0.003
						(0.074)	(0.076)	(0.019)	(0.008)
ln distance						-0.151^a	-0.127^a		
						(0.004)	(0.004)		
contiguity						0.108^a	0.095^a		
						(0.019)	(0.018)		
UN vote correlation						0.121^a	0.076^a	0.054^a	0.003
						(0.011)	(0.011)	(0.005)	(0.002)
sum of democracy indexes						0.069^a	0.053^a	0.024^a	0.011^a
						(0.007)	(0.006)	(0.003)	(0.001)
military alliance						0.081^a	0.127^a	0.110^a	0.027^a
						(0.011)	(0.010)	(0.007)	(0.003)
genetic distance						0.004	0.009^b		
						(0.004)	(0.003)		
multi. openness							-0.021^a	0.004^b	-0.001
							(0.004)	(0.002)	(0.001)
multi. open. × wars pre-1945							1.567^a	0.629^a	0.195^a
							(0.292)	(0.053)	(0.024)
Method	logit	logit	logit	LPM	LPM	LPM	LPM	cty pair FE	cty pair FE
Sample	pre-1945 pairs	whole	whole	whole	whole	whole	whole	whole	whole
Observations	1694	9836	9836	9836	9836	6395	6152	36701	35737
R^2	0.263	0.224	0.240	0.138	0.147	0.359	0.359	0.081	0.017

Our main variable of interest is the interaction term between old wars and RTA trade gains. Interaction terms have a non-straightforward interpretation in discrete choice models like the logit, because of their non linear nature (Ai and Norton, 2003). As explained in details above, we therefore resort to a linear probability model (LPM), which has the additional advantage of handling fixed effects more easily in our panel estimates. Column (4) is simply the LPM version of the logit specification of column (3). While this different estimation method naturally yields different coefficients, the signs and significance levels are preserved in column (4). Column (5) introduces the interaction term of trade gains with old wars. This interaction term enters positively and significant at the 1 percent level. This supports our hypothesis that economic gains and security gains are complement: Dyads with large estimated economic gains are more likely to enter a RTA, and this effect rises with the historic intensity of wars of the partners.

In column (6) we include a number of bilateral controls: The two most important gravity variables, namely geographical distance and contiguity, and a list of controls for political affinity (UN vote correlation, the sum of Polity IV reported democracy indices, a dummy for the existence of a military alliance and an index of genetic distance). All of those variables add to the likelihood of belonging to the same agreement as expected. To discriminate between the effect of probability vs realization of wars we also include the frequency of recent wars, which, according to our discussion in section 3.3.1, is expected to enter negatively through their deleterious effect on the political cost of negotiations. The coefficient is negative and significant at the 1 percent threshold. Comparing the coefficient of old war vs recent wars, the result suggests that a "window of opportunity" mechanism is at work. Having had a history of conflicts in the past makes a country-pair more likely to sign a RTA at the condition that their recent history is not too conflicting: Any exogenous event that prevents two ancient enemies to fight for some period improves the chances that they sign a RTA, with the consequence of reducing further the chances of conflict escalation. We quantify precisely the size of those effects later in the paper.

In spite of the inclusion of all these control variables and the resulting reduction by one third of the sample size, all the coefficients of interest in column (6) keep the expected sign and remain statistically significant at the 1 percent threshold. In particular the coefficient of economic gains is unaffected: This confirms that unobserved heterogeneity is already filtered out by the inclusion of $\hat{u}_{jit}$ in previous specifications. Regarding the coefficient of old wars, it is reduced significantly but it remains positive and significant.

Column (7) tests our implication 4, namely that multilateral trade openness and the probability of war have a positive and complementary impact on the RTA decision. As expected, the coefficient of the interaction term between multilateral openness and old war is positive; and it is highly significant.

Column (7) establishes our main results with a substantial set of controls, and we consider it as our benchmark specification. Remarkably the five coefficients of interest all have the expected sign and are statistically different from zero. According to our theoretical discussion in section 2.3, the fact that the coefficients of the two interaction terms are both positive confirms the internal consistency of our theoretical setup. As discussed in section 3.3 we can also infer from those coefficients that the point estimate of the theoretical elasticity $\hat{\varepsilon}_{cost}$ is 51.1. This estimate gives us an indication on the view of policy makers on the pacifying effect of RTAs: They believe that a 1 percent increase in costs of war divides by two the probability of war.

The two remaining columns extend the sample to the panel dimension. Both specifications include country-pair fixed effects. The coefficient on old wars cannot be estimated any more, but its interaction with trade gains can. For each dyad, we average data over non-overlapping time windows of 5 years, a method comparable to Egger and Larch (2008) and Martin et al. (2008) in related work. Column (8) considers the full sample. In column (9), we drop observations following the signature of RTA for those who do become members. This RTA onset specification is very demanding and, in spite of the five year averaging procedure, it is highly sensitive to measurement errors in the time-series dimension. With respect to the benchmark cross-sectional estimates in column (7), all the coefficients of interest keep their expected sign and are statistically significant, with the exception of the coefficient on new wars in the RTA onset specification. An important change is also the size of the coefficient on trade gains, when going from RTA (in col. 8) to RTA onset (in col.9) as a dependent variable. This was to be expected from our analysis of Figure 2 and from our discussion of the reverse causality issue: RTAs boost trade volumes, which reinforces even more the RTA-related economic gains after the signature.

Table 3 pushes further the robustness investigation. Those regressions take column (7) of Table 2 as a benchmark specification (with gravity controls unreported). In the first column, we re-estimate this benchmark specification using logit instead of LPM. It can be seen that all signs of the relevant variables are the same, that the global explanatory power is very high, and that the level of significance of the interaction term between old wars and trade gains is now slightly above ten percent (10.6% exactly). This logit estimate is the one which we use in the quantification section.

In the second column, we return to LPM and extend the set of gravity controls to include common language or legal system, colonial linkages, landlockness and remoteness of the country pair. All our variables of interest keep the same sign. Column (3) changes the definition of bilateral trade gains to be the average of the two countries RTA-related trade gains rather than the minimum. Given that the minimum is always smaller than the average, this translates mechanically into a decrease in the coefficient of trade gains.

Column (4) adds a set of dummy variables coding for each country, a feature which can be properly

identified in our cross-sectional sample of (non directional) country pairs. These dummy variables control for all time-invariant unobserved characteristics of a country that might make it more likely to fight wars in the past and to sign RTAs now. The global fit naturally increases substantially while leaving most of our results of interest remarkably similar. Column (5) adds a dummy to control for the fact that the two countries belong to the same geographical region of the world (following the World Bank definition of regions). This increases the probability of RTA significantly, while again leaving our results on trade gains and conflictuality unaffected.

Column (6) removes intra-EU observations by excluding all country-pairs where both countries belong to the European Union at 15. This is intended to check that our results are not entirely driven by European countries, which are characterized both by a rich history of warfares and by the creation of the worldwide deepest trade agreement. In this specification, all variables related to wars have slightly smaller coefficients, but they remain very significant.

Column (7) extends our definition of old wars by including a variable that accounts for war frequency 20 to 40 years before RTA signature. This results into a smoother representation of the history of wars with very recent ones, those that are more than one generation old, and the very old wars (before 1945). The pattern of coefficients is that recent wars tend to reduce the RTA probability, less recent ones tend to slightly promote them, while old wars have a much stronger positive effect. This finding matches well with our identification strategy: The political and subjective costs imposed by recent wars during RTA negotiation is gradually overturned by the positive strategic effect of war history.

4.2 Quantification and counterfactual experiments

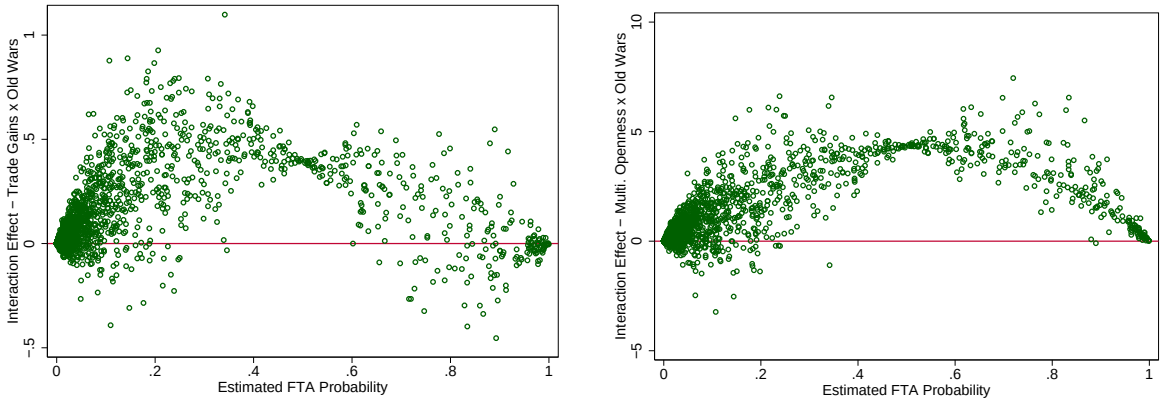
Up to this point, we have mostly analyzed the signs and statistical significance of coefficients. We now want to quantify the magnitude of the effects we have identified. In order to calculate counterfactuals we need to resort to a logit econometric model where the RTA probability cannot go outside the 0-1 range. Moreover the presence of interaction terms, which are the core of our analysis, are not straightforward in this context.

In all that follows we adopt the following strategy. We start by running a benchmark regression using logit (column 1 of Table 3), to estimate the coefficients of interest which gives us the benchmark probability of signing a RTA for each country pair in the sample. We then select a group of observations and we run a counterfactual by attributing them other values for one or more explanatory variables. For instance we take the country pairs in the lowest decile of the frequency of old war variable and we give them an artificial history of wars. Using the logit formula with the benchmark estimated coefficients, we recalculate their RTA probability and compare it with the benchmark probability to

Table 3: RTA determinants, robustness

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep var	RTA	RTA	RTA	RTA	RTA	RTA	RTA
war freq. pre-1945 (WAR_{ij})	44.866^a	8.209^a	4.637^a	6.175^a	6.075^a	3.823^a	6.046^a
	(15.989)	(0.754)	(0.900)	(0.670)	(0.662)	(0.676)	(0.662)
trade gains ($\hat{T}_{ijt}$)	0.296^a	0.007^a	0.003^b	0.005^a	0.005^a	0.004^a	0.005^a
	(0.042)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
trade gains × wars pre-1945	1.582	0.463^a	0.302^a	0.333^a	0.324^a	0.193^a	0.325^a
	(1.003)	(0.041)	(0.071)	(0.037)	(0.037)	(0.037)	(0.037)
war freq. $[t-20; t-1]$	-7.423^a	-0.464^a	-0.500^a	-0.188^a	-0.173^b	-0.154^b	-0.321^a
	(2.123)	(0.076)	(0.081)	(0.069)	(0.068)	(0.066)	(0.106)
multi. open. × wars pre-1945	17.364^a	1.684^a	0.777^b	1.446^a	1.396^a	0.865^a	1.375^a
	(5.980)	(0.291)	(0.325)	(0.257)	(0.254)	(0.253)	(0.254)
multi. openness	-1.995^a	-0.020^a	-0.027^a	-0.222^a	-0.216^a	-0.218^a	-0.217^a
	(0.233)	(0.004)	(0.005)	(0.012)	(0.012)	(0.011)	(0.012)
# of landlocked in dyad		-0.000	-0.004	-0.850^a	-0.855^a	0.044	-0.852^a
		(0.005)	(0.006)	(0.163)	(0.161)	(0.156)	(0.161)
common language		-0.019^b	-0.012	-0.020^a	-0.014^c	-0.012	-0.015^c
		(0.008)	(0.009)	(0.008)	(0.008)	(0.007)	(0.008)
colonial link		-0.031	-0.029	-0.075^a	-0.066^a	-0.052^a	-0.066^a
		(0.019)	(0.020)	(0.017)	(0.017)	(0.017)	(0.017)
common legal origin		-0.002	-0.010	-0.003	-0.017^a	-0.011^b	-0.017^a
		(0.006)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)
remoteness		0.083^a	0.097^a	-0.126^a	-0.153^a	-0.080^a	-0.152^a
		(0.011)	(0.013)	(0.025)	(0.025)	(0.024)	(0.025)
same region					0.114^a	0.050^a	0.116^a
					(0.010)	(0.010)	(0.010)
war freq. $[t-40; t-20]$							0.169^c
							(0.092)
Method	logit	LPM	LPM	cty FE	cty FE	cty FE	cty FE
Sample	whole	whole	whole	whole	whole	no EU15	whole
Trade Gains	min	min	average	min	min	min	min
Observations	6152	6152	5274	6152	6152	6071	6152
R^2	0.576	0.366	0.350	0.572	0.582	0.518	0.582

Figure 3: The interaction terms



(a) economic gains and security gains

(b) multi. openness and security gains

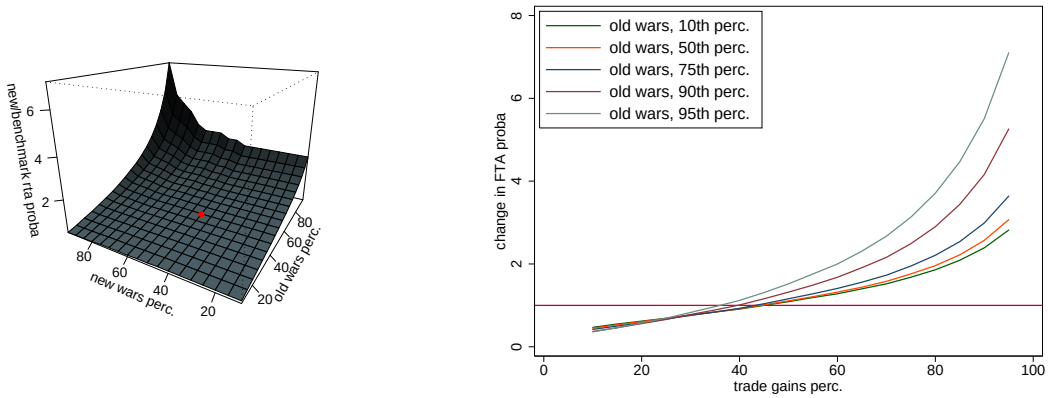
evaluate the magnitude of the effect of the altered variable. This procedure ensures that the probability remains in the admissible range, while doing a "what if" experiment: What if low conflict dyads had had an intense past history of warfare, keeping everything else constant?

4.2.1 Complementarity is a first-order effect

We first quantify our main result, that is the complementarity between old wars and trade gains in the formation of RTAs. The coefficient of the interaction term between trade gains and old wars is positive both in our benchmark LPM specification (col.7, Table 2) and in our benchmark logit specification (col. 1, Table 3). However Ai and Norton (2003) show that interaction terms have a sign that can be deceptive in a logit framework, and that cannot be interpreted readily. To investigate this question more fully, we calculate the marginal effect of this interaction term for the whole range of benchmark probabilities. In our case, we need to take care of the fact that old wars is interacted with trade gains, but also with multilateral openness. This complicates somewhat the computation of the marginal effects with respect to Ai and Norton (2003) and we report the details in the appendix. Results are graphed in figure 3. Each of the panels reports the marginal effects for each of the two interaction terms; each dot corresponds to an observed country-pair. We see that the marginal effects of the two interaction terms are very dominantly positive. It also confirms (see the computation in appendix) that, due to the functional form of the logit probability distribution, the reversal of the sign of the marginal effects is more likely when the estimated probabilities are located in the neighborhoods of 0 and 1. Since in our sample those estimated probabilities are quite concentrated at those two extreme values, verifying that those marginal effects are indeed positive was important.

We now turn to the quantification of the interaction term. To this purpose we choose pairs of

Figure 4: Complementarity between economic gains and security gains



(a) 3d view

(b) 2d view

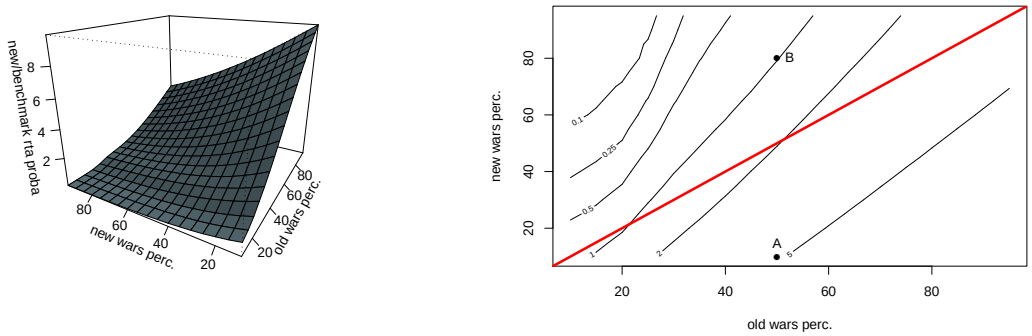
countries that are located inside the middle decile of those two variables, that is around the median level of old wars and trade gains. We then calculate the ratio of counterfactual to benchmark probabilities of RTA formation following the procedure just described, and spanning over the 10th to 95th percentiles of each variable. Results are in figure 4. In panel (a), it is clear that trade gains increase the probability of signing a RTA, and that the effect increases strongly with old wars. Panel (b) allows to better illustrate the effect. The x-axis reports trade gains while the y-axis reports the ratio of counterfactual to benchmark probabilities. Each curve corresponds to different levels of old wars. For a dyad that moves from the median to the top 20% of trade gains, the RTA probability is multiplied by two (1.96) if the dyad is in the middle range of old wars, while the multiplicative factor is almost 3 if the same dyad is in the top 10% of war history.[22] We see that the interaction term has a first-order importance. This confirms our intuition that trade gains are important mostly because they allow to increase security gains from RTA formation.

4.2.2 Windows of opportunity

Our second simulation uses the same method described at the start of this section to quantify the effect of the *probability* of war, measured by the frequency of old wars, and to compare it with the effect of the *realization* of war, measured by the frequency of recent wars. This is intended to highlight the existence of windows of opportunity during which interrupted conflict between old enemies may help sign a RTA and "lock in" a more peaceful bilateral relation. The left panel of figure 5 is very similar as the one in figure 4. We take the whole set of dyads with no history of recent nor old wars,

[22]The benchmark probabilities of signing a RTA in this precise sample have an average value of 7.7%. The median is much lower at 0.75%, which shows that most country pairs in the world have a very low RTA probability, while a few of them have a quite high one (ten percent of the sample has a benchmark probability higher than 20%).

Figure 5: Old wars and new wars



(a) 3dview

(b) 2d view

and gradually move them into the war space, looking at the changes in RTA probability. As expected from the point estimates in tables 2) and 3, recent wars reduce the probability of RTA formation, while old ones increase it. Again the magnitude of the effects is quite substantial. Panel (b) uncovers an interesting trade off that leaves the change in RTA probability unchanged. Panel (b) is a contour plot, where each curve represents a probability ratio from panel (a). Old wars are on the x-axis, recent wars on the y-axis. Assume a country goes from no old wars to the median level. This multiplies its benchmark RTA probability by almost five (point A in the figure) if there has been very little recent wars, while it leaves the probability unchanged if the level of recent wars is moved to the top 20% (point B in the figure). This shows that a change in old wars has in general a larger effect than a equivalent change in recent ones (as revealed by the 45 degree line in red). In other words, if a country-pair recent history of warfare perfectly reflects its long run history, then the net, overall effect of war is to increase the probability of RTA formation. By contrast, suppose now we assign the top 5% level of old wars to a country pair with no old wars. This multiplies by 10 its RTA probability if recent wars are very rare, but only by 3.5 if the country is also in top 5% of recent conflicts.

The former analysis was looking at pairs that did not experience any conflict in the real world. In figure 6 we take a totally different focus, and look at the effect of recent wars on country pairs that experienced a large set of conflicts in the recent years. We consider four different dyads, India-Pakistan, Greece-Turkey, Egypt-Israel, and Iran-Turkey. Out of those, Greece and Turkey are the only ones in a RTA (through the customs union signed between the EU and Turkey in 1996).[23] For those four pairs, our variable measuring the proportion of recent conflicts (over the last 20 sample

[23]The recent war frequency variable is therefore calculated for 1976-1996 for Greece-Turkey, and for 1980-2000 for the three other pairs.

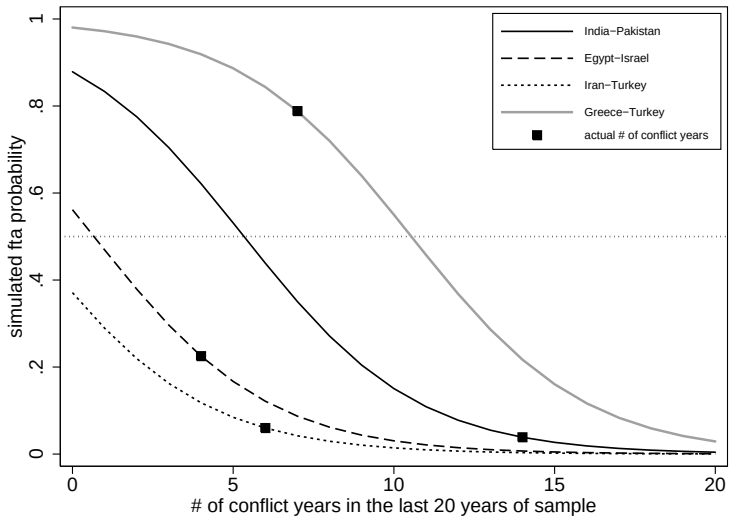


Figure 6: Window of opportunity for 4 emblematic country pairs

years) spans from 20 to 70% (4 to 14 years), with associated benchmark probability ranging from 4 to 80% as represented by the black squares on the graph. We then change the number of recent conflict years and calculate the new RTA probability. India-Pakistan is perhaps the most impressive example: After 5 years of peace, the RTA probability is multiplied by 5 at 20%, after ten years it jumps at 62%. Our results also reveal that 4 years of peace between Egypt and Israel brings their RTA probability from 23 to 57%. The effect of recent wars is quite abrupt for pairs that fundamentally have a large RTA signature probability (those with large potential economic gains, high proximity...). It thus suggests that the window of opportunity argument may be well grounded. For those pairs, even a short interruption of outbreaks in conflicts can increase RTA probability to a large extent and start a virtuous pacifying process. For Greece-Turkey, we observe the same overall shape of the impact of recent conflictuality, and note that in 1996, the conflictuality between the two countries seemed to have fallen to a level that made RTA possible.

4.2.3 A world without wars

Let us consider now the reverse counterfactual experiment. Instead of taking the peaceful dyads and make them fight, we make every country pair peaceful. That is the frequencies of old wars, recent wars and all their interaction terms are set to zero, and the resulting, counterfactual probabilities of RTA formation are estimated. Results are reported in figure 7, where the benchmark probability is on the x-axis, while the y-axis gives the counterfactual one (the dashed line corresponds to the 45 degree line). Each dot is a dyad, and some are singled out by symbols: Blue diamonds represent intra-EU pairs; red crosses represent country-pairs that were part of the communist bloc at some point; grey

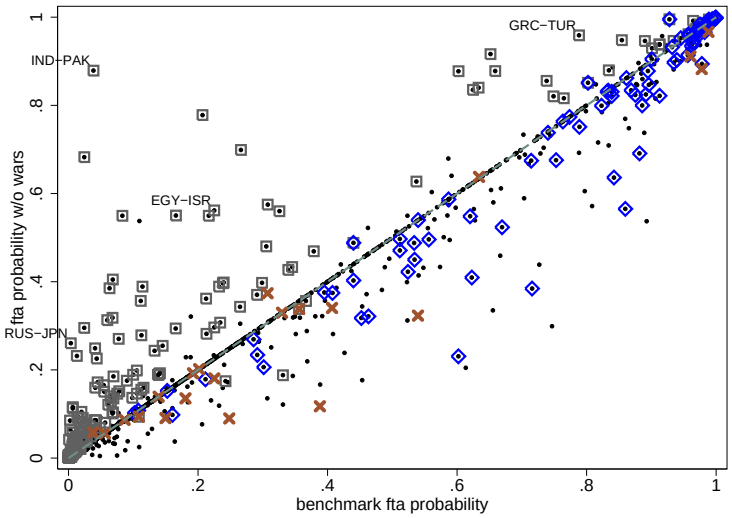


Figure 7: The world without military conflicts

squares represent pairs that have had a nonzero frequency of recent wars in the real world.

Both EU and former communist country pairs experience a drop in their counterfactual probability of RTA formation with respect to the benchmark one. This is, we believe, another illustration of the window of opportunity channel. Indeed, in both cases (both parts of the European continent in fact) the history of old wars is very intense. But the history of recent wars is very calm probably because the cold war made the two blocs very stable internally between the end of WWII and the collapse of the USSR. Those 45 years of "forced" peace between countries that used to fight a lot seems to have promoted the RTA wave in the region to a large extent.

Regarding the detrimental impact of recent wars, the examples of India-Pakistan and Egypt-Israel are probably the most illustrative. Those two pairs do have a very low level of benchmark probability of RTA formation; and this would jump to among the highest levels if one could cancel their history of recent wars. Greece-Turkey is another striking example.

4.2.4 Multilateralism triggers regionalism

In this quantification exercise, we look at the impact on RTA formation of multilateral trade openness. We simply cancel out multilateral globalization by setting multilateral trade openness to zero for all pairs of countries. Then we estimate the resulting, counterfactual probability of RTA formation that we compare to the benchmark probability. Results are reported in figure 8 where the grey triangles represent country pairs with an initial level of multilateral openness above the median level and where green diamonds represent pairs of countries belonging to Mercosur.

We see that in a counterfactual world without multilateral trade openness, most country pairs

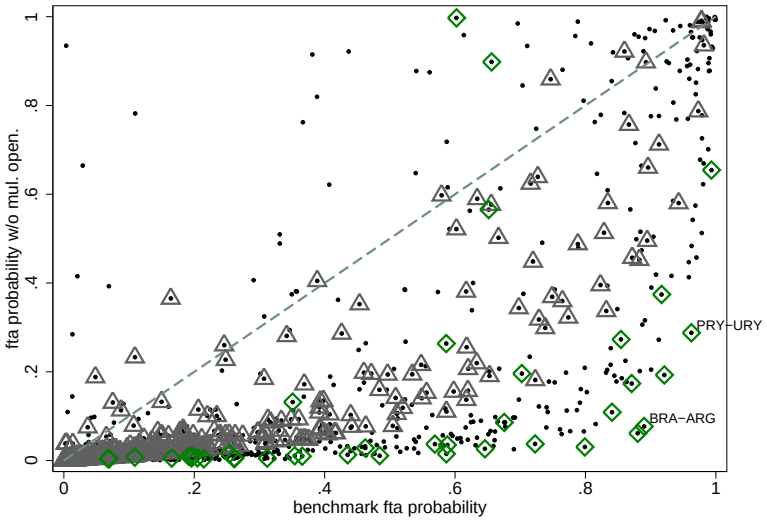


Figure 8: The world without multilateral trade

would experience a sharp decrease in their probability of RTA formation. This confirms our view (see section 2.3) that the wave of regionalism observed in the late 90s could be a policy response to the destabilizing, and conflict-promoting, effect of the development of multilateralism as experienced during the post world war period. This mechanism is particularly relevant for explaining the formation of Mercosur - a fact that has been widely discussed by policy practitioners (see Manzetti 1993).

5 Conclusion

Our results suggest that political scientists and historians are right to emphasize the political motivation behind RTAs, in particular the objective of pacifying relations. However, this does not mean that economics do not matter and that RTAs are signed without taking into account their economic benefits, the trade gains. On the contrary, without trade gains of RTAs that may be lost during a war, the peace promoting effect of RTAs is greatly weakened. Hence, our story is one where politics and economics push in the same direction. Economic and security gains are complementary to explain the evolving geography of trade agreements. Trade gains may be instrumentalized for a superior objective of peace but that makes them more, not less, important. Another important result is the interaction between multilateral and regional (or bilateral) trade liberalization. The recent multiplication of RTAs is often interpreted as a response of policy makers frustrated by stalling multilateral trade negotiations. Our result suggest a radically different story, one where multilateral openness (which may come from multilateral liberalization at WTO or the multiplication of RTAs) induces the formation of additional RTAs. RTAs can be interpreted as a way to reinforce bilateral economic relations between

countries at risk of war at a time when globalization reduces the bilateral economic dependence of these countries. The domino theory of regionalism of Baldwin (1995) comes to mind but here the danger that additional RTAs are attempting to counter is not the loss of economic attractiveness but the dangerous loss of economic dependency that it may imply. If X and Y (with an history of war) sign an RTA with Z (and more generally open to trade with the rest of the world), this induces them to sign a bilateral RTA between themselves to counter the relative loss of bilateral dependence this entails. Hence, RTAs may be contagious for political and not only for economic reasons. Finally, our results are consistent with the view that windows of opportunity for locking-in peace through trade exist. RTAs are difficult to sign for countries with an history of recent conflicts but country pairs with a long-run history of bilateral conflicts have a higher propensity to sign an RTA. Hence periods of peace between old enemies should be exploited to sign an RTA and lock-in a more peaceful bilateral relationship.

6 References

Baier S. and J. Bergstrand, 2004, Economic determinants of free trade agreements?, *Journal of International Economics* 64 29–63.

Baier S. and J. Bergstrand, 2007, Do free trade agreements actually increase members'international trade?, J*ournal of International Economics*, 71 (2007) 72–95

Baier S. and J. Bergstrand, 2008, Estimating the Effects of Free Trade Agreements on International Trade Flows using Matching Econometrics, *Journal of International Economics*, forthcoming.

Baldwin, R., 1995, "A Domino Theory of Regionalism," in *Expanding European Regionalism: The EU's New Members,* Baldwin, R.P. Haaparanta and J. Kiander (eds), Cambridge University Press.

Broda, C. and D. Weinstein, 2006 "Globalization and the Gains from Variety," *Quarterly Journal of Economics* 121(2):541–585.

Collier, P., Hoeffler, A., and Söderbom M., (2004), "On the Duration of Civil War", *Journal of Peace Research*, 41, 253-273.

Dimaranan, B.V., McDougall, R.A., 2002. Global Trade, Assistance, and Production: The GTAP 5 Data Base. Center for Global Trade Analysis, Purdue University.

Egger, P. and M., Larch, 2008, "Interdependent preferential trade agreement memberships: An empirical analysis", *Journal of International Economics*, vol.76.

Fearon, J. (1995), "Rationalist Explanations for War", *International Organization*, 49 (3), 379–414.

Fearon, J. (2005), "Primary Commodity Exports and Civil War", *Journal of Conflict Resolution*, 49, 483-507.

Fugazza, M. and F. Robert-Nicoud (2009), "The emulator effect of the Uruguay round on U.S. regionalism", mimeo UNCTAD.

Glick, R. and A.Taylor (2005), "Collateral damage: trade disruption and the economic impact of war", NBER WP 11565, forthcoming *Review of Economics and Statistics*.

Hertel, T. M. Ivanic, R. Keeney, 2007, "How Confident can we be in CGE-Based Assessments of Regional Trading Agreements?" *Economic Modeling* 24:611-635.

Keohan R. and J. Nye, 1977, Power and Interdependence, Boston, Little Brown.

Haas, E. 1980, "Why Collaborate? Issue linkage in international politics" *World Politics*, Vol. 32, No. 3 pp. 357-405

Imbens G. and J. Wooldridge, 2007, "Control Function and Related Methods" in *What's New in Econometrics?* NBER, Summer 2007, Lecture 6. `http://www.nber.org/WNE/lect_6_controlfuncs.pdf`

Levy, G. and Razin, R. (2004), "It Takes Two: An Explanation for the Democratic Peace", *Journal of the European Economic Association*, 2 (1), 1–29.

Mansfield E. and Jon C. Pevehouse, 2000, "Trade Blocs, Trade Flows, and International Conflict", *International Organization*, Vol. 54, No. 4, pp. 775-808

Manzetti, L., (1993), "The political economy of MERCOSUR", Journal of Interamerican Studies and World Affairs, 35 (4).

Martin, P., T. Mayer and M. Thoenig, 2008, "Make Trade not war?", *Review of Economic Studies* 75(3):865–900.

Martin, P., T. Mayer and M. Thoenig, 2006, *La mondialisation est-elle un facteur de paix ?* Paris : Ed. ENS rue d'Ulm, Opuscule # 3 du CEPREMAP. `http://www.cepremap.ens.fr/depot/opus/OPUS03.pdf`

Michaels G. and X. Zhi, 2007, "Freedom Fries", CEPR Discussion Paper 6380

Oneal, J. R. and B. M. Russett (1999b), "Assessing the Liberal Peace with Alternative Specifications: Trade Still Reduces Conflict", Journal of Peace Research, 36 (4): 423-442.

Polachek, S. (1980), "Conflict and Trade," Journal of Conflict Resolution, 24, 55-78.

Mocan, N., "Vengeance", 2008, NBER Working Paper No. 14131.

Powell, R. (1999) In the Shadow of Power: States and Strategies in International Politics (Princeton University Press).

Spoalaore Enrico and Romain Wacziarg, R. (2009) War and Relatdness, NBER Working Paper 15095.

Vicard, Vincent, 2008, Trade, conflicts and political integration : explaining the heterogeneity of regional trade agreements, mimeo.

7 Appendix

7.1 Theoretical Setup

In absence of RTA the expected welfare of a country is equal to $V = (1-\delta e)U_P + \delta e(1-W)U_P$. Indeed peace occurs with probability $(1-\delta e)$ and in that case the country trades under MFN tariff with its partner and gets the benchmark welfare U_P. War occurs with probability (δe); in that case trade is fully disrupted and some destructions happen; the country gets $(1-W)U_P$. If a RTA is in force the logic is similar and the expected welfare of the country is equal to $V^{RTA} = (1-\delta e^{RTA})(1+T)U_P + \delta e^{RTA}(1-W)U_P$. Plunging those two expressions into the RTA formation condition, $V^{RTA} - V \geq C$, and rescaling by the benchmark welfare U_P we easily get the condition (2).

Empirically we know that the probability of war is small: $\delta e \sim 1\%$. It is also likely that the cost of negotiation represents only a small fraction of the total welfare: $C/U_P \equiv c \sim 1\%$. We assume now that the RTA-related trade gains are small with respect to the cost of wars such that: $T/W \sim 1\%$. In spite of this reasonable assumption, we do not want to rule out the possibility for a RTA of having a first-order impact on the probability of escalation both through the opportunity cost channel and trough the political spillover channel. This means that we also have to assume in equation (3): $(\varepsilon_{pol}/\varepsilon_{cost}) \sim (T/W)$. Combining (2) and (3) we get:

$$(1-\delta e)T + (\delta e)W(1+\frac{T}{W})\frac{e^{RTA}-e}{e} \geq c \tag{16}$$

$$(1-\delta e)T + (\delta e)W\varepsilon_{cost}(1+\frac{T}{W})(\frac{\varepsilon_{pol}}{\varepsilon_{cost}}+\frac{T}{W}) \geq c \tag{17}$$

A first order approximation of this equation leads to

$$(1-\delta e)T + \delta e(\varepsilon_{pol}W + \varepsilon_{cost}T) \geq c \tag{18}$$

which corresponds to equation (3) in the main text.

Let consider the effect of multilateral openness. As discussed in the main text, multilateral openness corresponds to a decrease in the cost of war from W_0 to $W = (1-\omega)W_0$ and to an increase in the probability of escalation from e_0 to $e = (1+\varepsilon_{cost}\omega)e_0$. Substituting into equation (3) we get:

$$T + \varepsilon_{pol}\delta e_0 W_0 + (\varepsilon_{cost}-1)(\delta e_0 \times T) + (\varepsilon_{cost}-1)(\varepsilon_{pol}W_0 + \varepsilon_{cost}T)(\delta e_0 \times \omega) \geq c \tag{19}$$

In the previous equation the coefficient of the interaction term $(\delta e_0 \times \omega)$ is positive when $\varepsilon_{cost} > 1$. This proves the testable implication 4.

7.2 Marginal Effect and Interaction

Let denote x_1, x_2, x_3 our three variables of interest and $\mathbf{Z}$ the vector of covariates. Our Logit preferred specification (15) writes as

$$\widehat{P} = \frac{1}{1+\exp[-\beta_1 x_1 - \beta_2 x_2 - \beta_3 x_3 - \beta_{12}x_1x_2 - \beta_{13}x_1x_3 - \mathbf{Z^t}\beta]} \tag{20}$$

where $\widehat{P}$ is the estimated probability of FTA formation. Simple computations lead to

$$\frac{\partial^2 \widehat{P}}{\partial x_1 \partial x_2} = \widehat{P}(1-\widehat{P})\beta_{12} + \widehat{P}(1-\widehat{P})(1-2\widehat{P})(\beta_2 + \beta_{12}x_1)(\beta_1 + \beta_{12}x_2 + \beta_{13}x_3) \tag{21}$$

7.3 Further country pairs in trade gains table

Table 4: Estimated Trade Gains for the 51th-100th country-pairs in 1956

Country pair		Trade gains		bil. open.	dist.	ever fta?
		min T	mean T	min $\frac{\text{imports}}{\text{GDP}}$	kms	
USA	COL	.259%	2.262%	.105%	4251	No
FRA	**ITA**	**.256%**	**.338%**	**.261%**	**892**	**Yes**
THA	IDN	.256%	.305%	.615%	2306	Yes
GBR	DEU	.256%	.29%	.526%	809	Yes
NLD	IDN	.249%	.548%	1.363%	11346	No
CAN	VEN	.248%	.262%	.683%	4647	No
BEL	SWE	.244%	.3%	.941%	1152	Yes
DEU	DNK	.239%	.845%	.425%	538	Yes
CZS	BGR	.236%	.72%	.307%	1084	No
GBR	ZAF	.234%	1.474%	.481%	9489	Yes
JPN	IDN	.23%	.731%	.329%	5482	No
FRA	SWE	.227%	.262%	.231%	1616	Yes
HUN	ROM	.225%	.235%	.128%	540	Yes
SAU	IND	.223%	.3%	.191%	3509	No
DEU	AUT	.222%	1.358%	.393%	592	Yes
CHN	LKA	.219%	.426%	.095%	4914	No
CHN	JPN	.216%	.3%	.167%	1975	No
FRA	CHE	.214%	.57%	.217%	474	Yes
ARG	GBR	.213%	.3%	.438%	11137	No
CZS	ROM	.212%	.456%	.275%	902	No
HUN	BGR	.211%	.256%	.278%	693	Yes
GBR	IRL	.209%	2.164%	.429%	425	Yes
BRA	SWE	.209%	.247%	.545%	10185	No
IND	PAK	.208%	.271%	.178%	1238	No
VEN	NLD	.207%	.226%	.571%	7972	No
POL	AUT	.203%	.211%	.227%	549	Yes
BGR	ROM	.2%	.257%	.105%	370	Yes
BRA	URY	.193%	.568%	.368%	2168	Yes
SDN	EGY	.193%	.644%	.462%	1736	Yes
USA	BEL	.192%	.601%	.078%	7303	No
BRA	DNK	.19%	.191%	.365%	9776	No
ROM	EGY	.19%	.198%	.1%	1792	No
POL	HUN	.19%	.298%	.196%	520	Yes
ARG	ITA	.188%	.298%	.281%	11214	No
CHL	ARG	.184%	.338%	.255%	1157	Yes
SYR	SAU	.182%	.278%	.686%	1463	No
BRA	FIN	.179%	.219%	.34%	10749	No
HUN	CHN	.178%	.336%	.077%	7710	No
GBR	BEL	.177%	.417%	.363%	448	Yes
IDN	AUS	.175%	.211%	.501%	5078	No
CHE	AUT	.175%	.277%	.587%	576	Yes
ARG	DEU	.174%	.473%	.309%	11646	No
BRA	ESP	.174%	.201%	.206%	7821	No
LBN	SAU	.172%	.226%	.648%	1417	No
HND	SLV	.171%	.274%	.519%	244	Yes
JPN	AUS	.171%	.404%	.346%	7827	No
SYR	JOR	.169%	.317%	.733%	373	No
BRA	NOR	.168%	.169%	.324%	10018	No
AFG	PAK	.168%	.257%	.104%	806	No
ITA	SWE	.166%	.194%	.248%	1833	Yes

Note: Lines in boldface indicate pairs that sign the Rome Treaty establishing the European Economic Community a year later.